INCAST 2008-124

A STUDY ON INTERCONNECTIVITY FOR FLOSOLVER TERAFLOP PARALLEL MACHINE

G.Sundara Rao[1] and T.Venkatesh[2]

[1] Flosolver Unit, National Aerospace Laboratories,Bangalore,India, sundar@flomail.nal.res.in
[2] Department of computer Science & Engineering, Ghousia College of Engineering, Ramanagaram, India, t_venk5@yahoo.com

ABSTRACT: India's first parallel computer, Flosolver Mark1 (Mk1), was developed by Flosolver Unit, National Aerospace Laboratories (NAL) in 1986[1]. Since then seven versions of parallel computers have been developed using the state-of-the-art processors and presently engaged in building the eighth version, a Tera flop parallel machine. In parallel machines connectivity plays a major role. In the year 2000 a communication device, FloSwitch, was indigenously designed, developed and used in the fifth version, Flosolver Mk5[2]. All the later developed machines are also based on this device or its updated versions. FloSwitch has its unique features like parallel read/write, broadcast write and it supports both message passing and message processing mechanism. In this paper it is about one of the possible connectivity, routing. The interconnectivity was proposed with respect to the indigenously developed FloSwitch and optical Module[2]. Flosolver uses the cluster-based architecture, Fig.1, where each cluster contains 4 nodes, each with dual processors and a PCI-DPM card[3] and these nodes are connected to the FloSwitch. FloSwitch handles the communication within the clusters and across the clusters communication will be through optical module.

1. INTRODUCTION

Basic cluster architecture: A cluster is a type of parallel/distributed processing architecture consisting of a set of interconnected computers that can work as a single machine[4]. Here all the clusters are homogeneous.

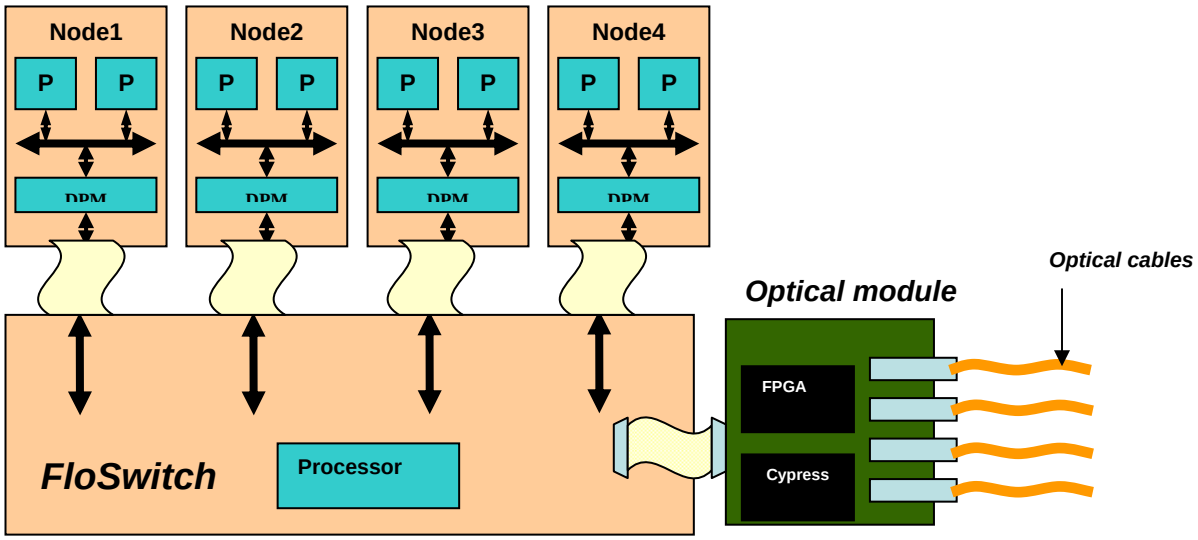

Fig.1 Flosolver cluster: A schematic diagram showing connectivity between Nodes, FloSwitch and optical module.

2. FLOSOLVER INTERCONNECTIVITY

Our present requirement is to communicate among 1024 processors placed in 128 clusters, each cluster having 16 optical links. If all to all connection is made then it requires $(128)^2$ connections which is complex. So in order to balance between optimal number of hops and optical link connections Base 4 Connectivity is discussed here.

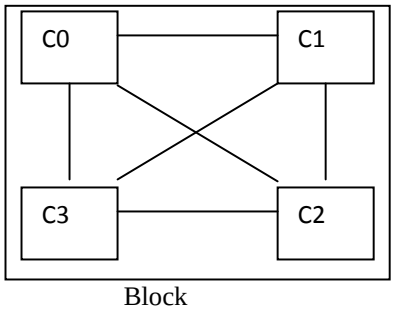

Fig.2a: optical interconnection between four clusters

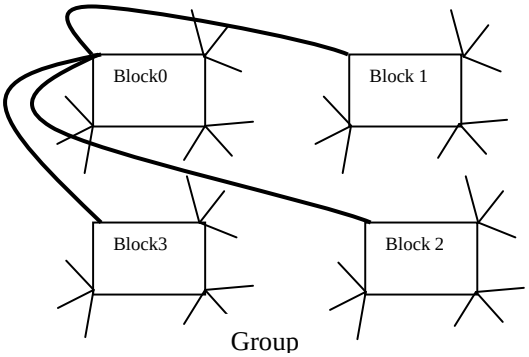

Fig.2b: Schematic diagram showing the connections from one cluster of a Block. (other connections not shown for clarity)

Fig.2a represents basic connectivity of the four clusters (Block). In the next level four such Blocks (B0, B1, B2, B3) are connected (Fig 2b). Using this method, first level requires 1 hop and second level requires 2 hops. Also every switch will go in factors of 4 and it requires atmost 3 more optical links. Four such blocks are connected to form a group, and four such groups (G0, G1, G2, G3) are connected to form one region. Similarly another set of region will be formed, and these two regions (R0, R1) are connected through optical link. Each cluster will be having sixteen optical links. In this method we are using ten optical links to connect the 1024 processors. While the Figure 3 shows the representation of complete machine, Figure 4 shows its tree representation the maximum hops will be 4. Labeling of the clusters is also shown in Figure 4, where the unit place represents cluster number, next higher bit represents block number, next higher order bit represents group number and MSB represents the region number.

A methodology has been derived for calculating the number of hops required to reach from one cluster to another cluster based on the rule, where if one digit of a particular cluster is varied and other three digits of the cluster remains same with the other cluster which need to be interconnected then the number of hops required will be 1. Similarly if two digits are varied and two digits remain same with another cluster then two hops are needed to interconnect. Table 1shows sample examples. But at every stage three digits remains same and only one digit need to be changed to reach the destination.

Table 1: Hops requirement in Tera flop parallel machine

Source Cluster	**Destination Cluster**	**Hop**
0000	0001	1
0000	0203	2
0000	0233	3
0123	0323	1
0122	0333	3
0000	1222	4
0333	1333	1
0000	1000	1
0000	1001	2
0131	1023	4

Example:

hop--> 1 2 3 4 [shown in black dotted lines Fig3]

1) (source) 0000 --------->1000 --------->1200--------->1220--------->1222 (destination)

hop--> 1 2 3

2) (source) 0122--------->0123---------->0323---------->0333(destination)

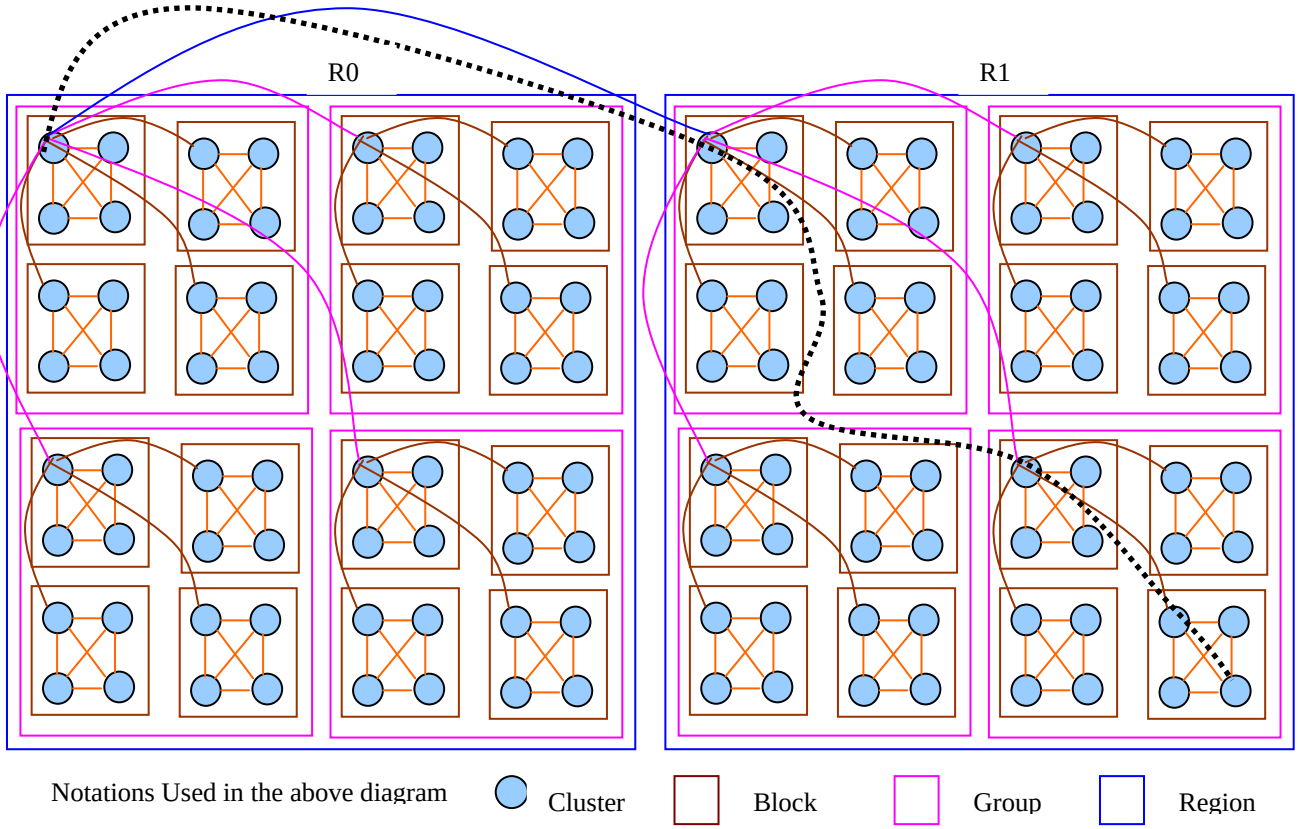

Fig 3: complete system showing the clusters, blocks, groups and Regions. For simplicity connections from one cluster only are drawn among blocks, groups, and regions. Similar connections have to be drawn from other clusters.

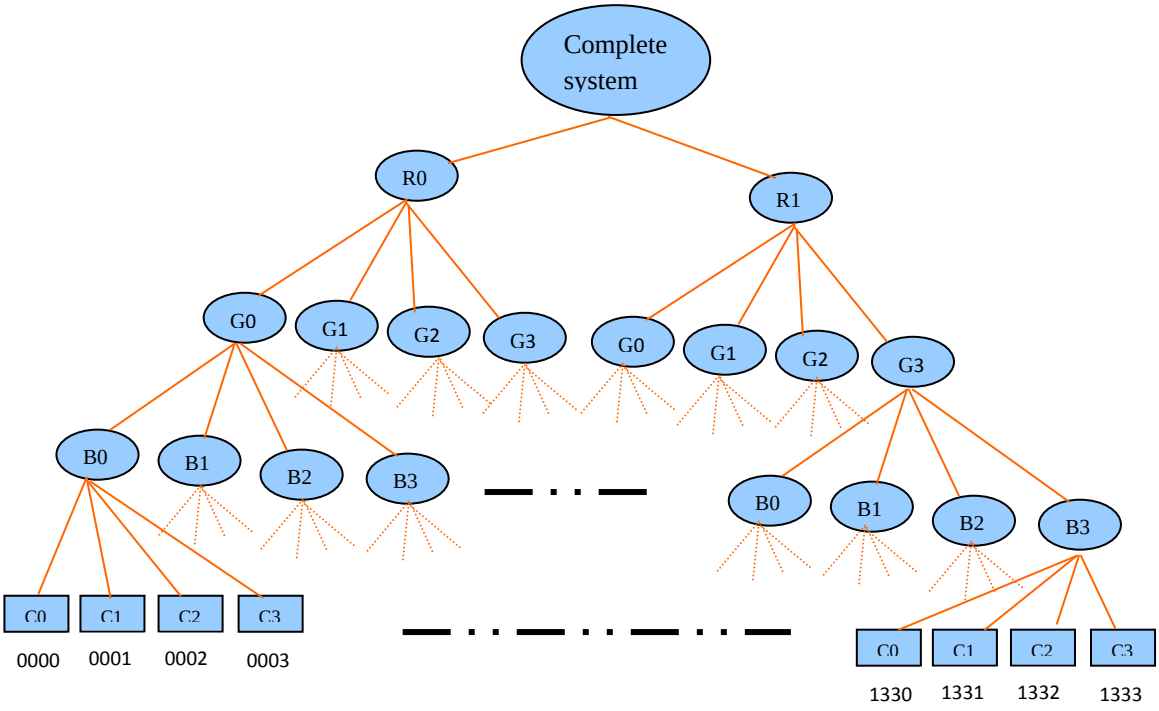

Fig 4:Tree representation to show numbering

3. CONCLUSION AND FUTURE WORK

In this paper base 4 connectivity has been discussed. This is one of the possible connectivity. This method uses Ten optical links and maximum hops required will be four. Studies are going on to workout other topologies to connect the 1024 processor system and to reduce the hops to three. This method ensures the simultaneous communication between nodes and clusters without any bottleneck. By this method we can connect 2048 processor using twelve optical links, which also requires maximum of four hops.

4. ACKNOWLEDGMENTS

We wish to acknowledge Dr.A.R.Upadhaya, Director, NAL for his encouragement and support. We are sincerely thankful to Dr. U.N.Sinha, Head, Flosolver Unit, NAL, Bangalore for his guidance and encouragement during the course of investigation.

5. REFERENCES

[1] U.N.Sinha, M.D Deshpande and V.R Sarasamma, “Flosolver: A Parallel Computer for Fluid Dynamics”. Current Science, Vol. 57, pp. 1277-1285, 1988.

[2] Aravinda K M, Kalyani Sinha, Charles Niranjan, Lakshman L, Arun Kumar S, Rakesh Kumar Jha, Ananda Raj D, Chandra Kumar C, Jaganadham VV and Sunadar Rao G. “Design of i486 based Flo-Opti-link Switch”. NAL PDFS 0305, Feb 2003.

[3] Flosolver Team, “Line diagrams of PCI-DPM Card”, NAL PD FS 0016, Nov 2000.

[4] Pfister G. “In search of clusters”, Prentice Hall, 2nd edition, 1998.